ATGAAATTTA GTAAAAAATA TATAGCAGCT GGATCAGCTG TTATCGTATC CTTGAGTCTA 60
TGTGCCTATG CACTAAACCA GCATCGTTCG CAGGAAAATA AGGACAATAA TCGTGTCTCT 120
TATGTGGATG GCAGCCAGTC AAGTCAGAAA AGTGAAAACT TGACACCAGA CCAGGTTAGC 180
CAGAAAGAAG GAATTCAGGC TGAGCAAATT GTAATCAAAA TTACAGATCA GGGCTATGTA 240
ACGTCACACG GTGACCACTA TCATTACTAT AATGGGAAAG TTCCTTATGA TGCCCTCTTT 300
AGTGAAGAAC TCTTGATGAA GGATCCAAAC TATCAACTTA AAGACGCTGA TATTGTCAAT 360
GAAGTCAAGG GTGGTTATAT CATCAAGGTC GATGGAAAAT ATTATGTCTA CCTGAAAGAT 420
GCAGCTCATG CTGATAATGT TCGAACTAAA GATGAAATCA ATCGTCAAAA ACAAGAACAT 480
GTCAAAGATA ATGAGAAGGT TAACTCTAAT GTTGCTGTAG CAAGGTCTCA GGGACGATAT 540
ACGACAAATG ATGGTTATGT CTTTAATCCA GCTGATATTA TCGAAGATAC GGGTAATGCT 600
TATATCGTTC CTCATGGAGG TCACTATCAC TACATTCCCA AAAGCGATTT ATCTGCTAGT 660
GAATTAGCAG CAGCTAAAGC ACATCTGGCT GGAAAAAATA TGCAACCGAG TCAGTTAAGC 720
TATTCTTCAA CAGCTAGTGA CAATAACACG CAATCTGTAG CAAAAGGATC AACTAGCAAG 780
CCAGCAAATA AATCTGAAAA TCTCCAGAGT CTTTTGAAGG AACTCTATGA TTCACCTAGC 840
GCCCAACGTT ACAGTGAATC AGATGGCCTG GTCTTTGACC CTGCTAAGAT TATCAGTCGT 900
ACACCAAATG GAGTTGCGAT TCCGCATGGC GACCATTACC ACTTTATTCC TTACAGCAAG 960
CTTTCTGCTT TAGAAGAAAA GATTGCCAGA ATGGTGCCTA TCAGTGGAAC TGGTTCTACA 1020
GTTTCTACAA ATGCAAAACC TAATGAAGTA GTGTCTAGTC TAGGCAGTCT TTCAAGCAAT 1080
CCTTCTTCTT TAACGACAAG TAAGGAGCTC TCTTCAGCAT CTGATGGTTA TATTTTTAAT 1140
CCAAAAGATA TCGTTGAAGA AACGGCTACA GCTTATATTG TAAGACATGG TGATCATTTC 1200
CATTACATTC CAAAATCAAA TCAAATTGGG CAACCGACTC TTCCAAACAA TAGTCTAGCA 1260
ACACCTTCTC CATCTCTTCC AATCAATCCA GGAACTTCAC ATGAGAAACA TGAAGAAGAT 1320
GGATACGGAT TTGATGCTAA TCGTATTATC GCTGAAGATG AATCAGGTTT TGTCATGAGT 1380
CACGGAGACC ACAATCATTA TTTCTTCAAG AAGGACTTGA CAGAAGAGCA AATTAAGGCT 1440
GCGCAAAAAC ATTTAGAGGA AGTTAAAACT AGTCATAATG GATTAGATTC TTTGTCATCT 1500
CATGAACAGG ATTATCCAGG TAATGCCAAA GAAATGAAAG ATTTAGATAA AAAAATCGAA 1560
GAAAAAATTG CTGGCATTAT GAAACAATAT GGTGTCAAAC GTGAAAGTAT TGTCGTGAAT 1620
AAAGAAAAAA ATGCGATTAT TTATCCGCAT GGAGATCACC ATCATGCAGA TCCGATTGAT 1680
GAACATAAAC CGGTTGGAAT TGGTCATTCT CACAGTAACT ATGAACTGTT TAAACCCGAA 1740
GAAGGAGTTG CTAAAAAAGA AGGGAATAAA GTTTATACTG GAGAAGAATT AACGAATGTT 1800
GTTAATTTGT TAAAAAATAG TACGTTTAAT AATCAAAACT TTACTCTAGC CAATGGTCAA 1860
AAACGCGTTT CTTTTAGTTT TCCGCCTGAA TTGGAGAAAA AATTAGGTAT CAATATGCTA 1920
GTAAAATTAA TAACACCAGA TGGAAAAGTA TTGGAGAAAG TATCTGGTAA AGTATTTGGA 1980
GAAGGAGTAG GGAATATTGC AAACTTTGAA TTAGATCAAC CTTATTTACC AGGACAAACA 2040
TTTAAGTATA CTATCGCTTC AAAAGATTAT CCAGAAGTAA GTTATGATGG TACATTTACA 2100
GTTCCAACCT CTTTAGCTTA CAAAATGGCC AGTCAAACGA TTTTCTATCC TTTCCATGCA 2160
GGGGATACTT ATTTAAGAGT GAACCCTCAA TTTGCAGTGC CTAAAGGAAC TGATGCTTTA 2220
GTCAGAGTGT TTGATGAATT TCATGGAAAT GCTTATTTAG AAAATAACTA TAAAGTTGGT 2280
GAAATCAAAT TACCGATTCC GAAATTAAAC CAAGGAACAA CCAGAACGGC CGGAAATAAA 2340
ATTCCTGTAA CCTTCATGGC AAATGCTTAT TTGGACAATC AATCGACTTA TATTGTGGAA 2400
GTACCTATCT TGGAAAAAGA AAATCAAACT GATAAACCAA GTATTCTACC ACAATTTAAA 2460
AGGAATAAAG CACAAGAAAA CTCAAAACTT GATGAAAAGG TAGAAGAACC AAAGACTAGT 2520
GAGAAGGTAG AAAAAGAAAA ACTTTCTGAA ACTGGGAATA GTACTAGTAA TTCAACGTTA 2580
GAAGAAGTTC CTACAGTGGA TCCTGTACAA GAAAAAGTAG CAAAATTTGC TGAAAGTTAT 2640
GGGATGAAGC TAGAAAATGT CTTGTTTAAT ATGGACGGAA CAATTGAATT ATATTTACCA 2700

TCAGGAGAAG TCATTAAAAA GAATATGGCA GATTTTACAG GAGAAGCACC TCAAGGAAAT 2760
GGTGAAAATA AACCATCTGA AAATGGAAAA GTATCTACTG GAACAGTTGA GAACCAACCA 2820
ACAGAAAATA AACCAGCAGA TTCTTTACCA GAGGCACCAA ACGAAAAACC TGTAAAACCA 2880
GAAAACTCAA CGGATAATGG AATGTTGAAT CCAGAAGGGA ATGTGGGGAG TGACCCTATG 2940
TTAGATCCAG CATTAGAGGA AGCTCCAGCA GTAGATCCTG TACAAGAAAA ATTAGAAAAA 3000
TTTACAGCTA GTTACGGATT AGGCTTAGAT AGTGTTATAT TCAATATGGA TGGAACGATT 3060
GAATTAAGAT TGCCAAGTGG AGAAGTGATA AAAAAGAATT TATCTGATTT CATAGCGTAA 3120
(SEQ ID NO: 1)
FIGURE 1

AATTCCTTGT CGGGTAAGTT CCGACCCGCA CGAAAGGCGT AATGATTTGG GCACTGTCTC 60
AACGAGAGAC TCGGTGAAAT TTTAGTACCT GTGAAGATGC AGGTTACCCG CGACAGGACG 120
GAAAGACCCC ATGGAGCTTT ACTGCAGTTT GATATTGAGT GTCTGTACCA CATGTACAGG 180
ATAGGTAGGA GTCTAAGAGA TCGGGACGCC AGTTTCGAAG GAGACGCTGT TGGGATACTA 240
CCCTTGTGTT ATGGCCACTC TAACCCAGAT AGGTGATCCC TATCGGAGAC AGTGTCTGAC 300
GGGCAGTTTG ACTGGGGCGG TCGCCTCCTA AAAGGTAACG GAGGCGCCCA AAGGTTCCCT 360
CAGAATGGTT GGAAATCATT CGCAGAGTGT AAAGGTATAA GGGAGCTTGA CTGCGAGAGC 420
TACAACTCGA GCAGGGACGA AAGTCGGGCT TAGTGATCCG GTGGTTCCGT ATGGAAGGGC 480
CATCGCTCAA CGGATAAAAG CTACCCTGGG GATAACAGGC TTATCTCCCC CAAGAGTTCA 540
CATCGACGGG GAGGTTTGGC ACCTCGATGT CGGCTCGTCG CATCCTGGGG CTGTAGTCGG 600
TCCCAAGGGT TGGGCTGTTC GCCCATTAAA GCGGCACGCG AGCTGGGTTC AGAACGTCGT 660
GAGACAGTTC GGTCCCTATC CGTCGCGGGC GTAGGAAATT TGAGAGGATC TGCTCCTAGT 720
ACGAGAGGAC CAGAGTGGAC TTACCGCTGG TGTACCAGTT GTCTTGCCAA AGGCATCGCT 780
GGGTAGCTAT GTAGGGAAGG GATAAACGCT GAAAGCATCT AAGTGTGAAA CCCACCTCAA 840
GATGAGATTT CCCATGATTA TATATCAGTA AGAGCCCTGA GAGATGATCA GGTAGATAGG 900
TTAGAAGTGG AAGTGTGGCG ACACATGTAG CGGACTAATA CTAATAGCTC GAGGACTTAT 960
CCAAAGTAAC TGAGAATATG AAAGCGAACG GTTTTCTTAA ATTGAATAGA TATTCAATTT 1020
TGAGTAGGTA TTACTCAGAG TTAAGTGACG ATAGCCTAGG AGATACACCT GTACCCATGC 1080
CGAACACAGA AGTTAAGCCC TAGAACGCCG GAAGTAGTTG GGGGTTGCCC CCTGTGAGAT 1140
AGGGAAGTCG CTTAGCTCTA GGGAGTTTAG CTCAGCTGGG AGAGCATCTG CCTTACAAGC 1200
AGAGGGTCAG CGGTTCGATC CCGTTAACTC CCAAAGGTCC CGTAGTGTAG CGGTTATCAC 1260
GTCGCCCTGT CACGGCGAAG ATCGCGGGTT CGATTCCCGT CGGGACCGTT TAAGGTAACG 1320
CAAGTTATTT TAGACTCGTT AGCTCAGTTG GTAGAGCAAT TGACTTTTAA TCAATGGGTC 1380
ACTGGTTCGA GCCCAGTACG GGTCATATAT GCGGGTTTGG CGGAATTCTA ATCTCTTTGA 1440
AATCATCTTC TCTCACTTTC CAAAACTCTA TTACCTCTTA TTATACCACA TTTCAATCTT 1500
CAACTTCCCA GTAATATAAG CACCTCTGGC GAAAGAAGTT TCAATGTCCT AAAGTAATAA 1560
GTGAATCCAA TTCAGGAACT CCAAGAACAA AAGAAACATC TGGTGTCACA AGTATTGGAT 1620
GGCACAGAGT CACGTGGTAG TCTGACCCTA GCAGAAATTT TAAATAGTAA ACTATTTACT 1680
GGTTAATTAA ATGGTTAAAT AACCGGTTTA GAAAACTATT TAATAAAGTA AAAGAAGTTG 1740
AGAAAAAACT TCATCATTTA TTGAAATGAG GGATTTATGA AATTTAGTAA AAAATATATA 1800
GCAGCTGGAT CAGCTGTTAT CGTATCCTTG AGTCTATGTG CCTATGCACT AAACCAGCAT 1860
CGTTCGCAGG AAAATAAGGA CAATAATCGT GTCTCTTATG TGGATGGCAG CCAGTCAAGT 1920
CAGAAAAGTG AAAACTTGAC ACCAGACCAG GTTAGCCAGA AAGAAGGAAT TCAGGCTGAG 1980
CAAATTGTAA TCAAAATTAC AGATCAGGGC TATGTAACGT CACACGGTGA CCACTATCAT 2040

TACTATAATG GGAAAGTTCC TTATGATGCC CTCTTTAGTG AAGAACTCTT GATGAAGGAT 2100
CCAAACTATC AACTTAAAGA CGCTGATATT GTCAATGAAG TCAAGGGTGG TTATATCATC 2160
AAGGTCGATG GAAAATATTA TGTCTACCTG AAAGATGCAG CTCATGCTGA TAATGTTCGA 2220
ACTAAAGATG AAATCAATCG TCAAAAACAA GAACATGTCA AAGATAATGA GAAGGTTAAC 2280
TCTAATGTTG CTGTAGCAAG GTCTCAGGGA CGATATACGA CAAATGATGG TTATGTCTTT 2340
AATCCAGCTG ATATTATCGA AGATACGGGT AATGCTTATA TCGTTCCTCA TGGAGGTCAC 2400
TATCACTACA TTCCCAAAAG CGATTTATCT GCTAGTGAAT TAGCAGCAGC TAAAGCACAT 2460
CTGGCTGGAA AAAATATGCA ACCGAGTCAG TTAAGCTATT CTTCAACAGC TAGTGACAAT 2520
AACACGCAAT CTGTAGCAAA AGGATCAACT AGCAAGCCAG CAAATAAATC TGAAAATCTC 2580
CAGAGTCTTT TGAAGGAACT CTATGATTCA CCTAGCGCCC AACGTTACAG TGAATCAGAT 2640
GGCCTGGTCT TTGACCCTGC TAAGATTATC AGTCGTACAC CAAATGGAGT TGCGATTCCG 2700
CATGGCGACC ATTACCACTT TATTCCTTAC AGCAAGCTTT CTGCTTTAGA AGAAAAGATT 2760
GCCAGAATGG TGCCTATCAG TGGAACTGGT TCTACAGTTT CTACAAATGC AAAACCTAAT 2820
GAAGTAGTGT CTAGTCTAGG CAGTCTTTCA AGCAATCCTT CTTCTTTAAC GACAAGTAAG 2880
GAGCTCTCTT CAGCATCTGA TGGTTATATT TTTAATCCAA AAGATATCGT TGAAGAAACG 2940
GCTACAGCTT ATATTGTAAG ACATGGTGAT CATTTCCATT ACATTCCAAA ATCAAATCAA 3000
ATTGGGCAAC CGACTCTTCC AAACAATAGT CTAGCAACAC CTTCTCCATC TCTTCCAATC 3060
AATCCAGGAA CTTCACATGA GAAACATGAA GAAGATGGAT ACGGATTTGA TGCTAATCGT 3120
ATTATCGCTG AAGATGAATC AGGTTTTGTC ATGAGTCACG GAGACCACAA TCATTATTTC 3180
TTCAAGAAGG ACTTGACAGA AGAGCAAATT AAGGCTGCGC AAAAACATTT AGAGGAAGTT 3240
AAAACTAGTC ATAATGGATT AGATTCTTTG TCATCTCATG AACAGGATTA TCCAGGTAAT 3300
GCCAAAGAAA TGAAAGATTT AGATAAAAAA ATCGAAGAAA AAATTGCTGG CATTATGAAA 3360
CAATATGGTG TCAAACGTGA AAGTATTGTC GTGAATAAAG AAAAAAATGC GATTATTTAT 3420
CCGCATGGAG ATCACCATCA TGCAGATCCG ATTGATGAAC ATAAACCGGT TGGAATTGGT 3480
CATTCTCACA GTAACTATGA ACTGTTTAAA CCCGAAGAAG GAGTTGCTAA AAAAGAAGGG 3540
AATAAAGTTT ATACTGGAGA AGAATTAACG AATGTTGTTA ATTTGTTAAA AAATAGTACG 3600
TTTAATAATC AAAACTTTAC TCTAGCCAAT GGTCAAAAAC GCGTTTCTTT TAGTTTTCCG 3660
CCTGAATTGG AGAAAAAATT AGGTATCAAT ATGCTAGTAA AATTAATAAC ACCAGATGGA 3720
AAAGTATTGG AGAAAGTATC TGGTAAAGTA TTTGGAGAAG GAGTAGGGAA TATTGCAAAC 3780
TTTGAATTAG ATCAACCTTA TTTACCAGGA CAAACATTTA AGTATACTAT CGCTTCAAAA 3840
GATTATCCAG AAGTAAGTTA TGATGGTACA TTTACAGTTC CAACCTCTTT AGCTTACAAA 3900
ATGGCCAGTC AAACGATTTT CTATCCTTTC CATGCAGGGG ATACTTATTT AAGAGTGAAC 3960
CCTCAATTTG CAGTGCCTAA AGGAACTGAT GCTTTAGTCA GAGTGTTTGA TGAATTTCAT 4020
GGAAATGCTT ATTTAGAAAA TAACTATAAA GTTGGTGAAA TCAAATTACC GATTCCGAAA 4080
TTAAACCAAG GAACAACCAG AACGGCCGGA AATAAAATTC CTGTAACCTT CATGGCAAAT 4140
GCTTATTTGG ACAATCAATC GACTTATATT GTGGAAGTAC CTATCTTGGA AAAAGAAAAT 4200
CAAACTGATA AACCAAGTAT TCTACCACAA TTTAAAAGGA ATAAAGCACA AGAAAACTCA 4260
AAACTTGATG AAAAGGTAGA AGAACCAAAG ACTAGTGAGA AGGTAGAAAA AGAAAAACTT 4320
TCTGAAACTG GGAATAGTAC TAGTAATTCA ACGTTAGAAG AAGTTCCTAC AGTGGATCCT 4380
GTACAAGAAA AAGTAGCAAA ATTTGCTGAA AGTTATGGGA TGAAGCTAGA AAATGTCTTG 4440
TTTAATATGG ACGGAACAAT TGAATTATAT TTACCATCAG GAGAAGTCAT TAAAAAGAAT 4500
ATGGCAGATT TTACAGGAGA AGCACCTCAA GGAAATGGTG AAAATAAACC ATCTGAAAAT 4560
GGAAAAGTAT CTACTGGAAC AGTTGAGAAC CAACCAACAG AAAATAAACC AGCAGATTCT 4620
TTACCAGAGG CACCAAACGA AAAACCTGTA AAACCAGAAA ACTCAACGGA TAATGGAATG 4680
TTGAATCCAG AAGGGAATGT GGGGAGTGAC CCTATGTTAG ATCCAGCATT AGAGGAAGCT 4740

```
CCAGCAGTAG ATCCTGTACA AGAAAAATTA GAAAAATTTA CAGCTAGTTA CGGATTAGGC     4800
TTAGATAGTG TTATATTCAA TATGGATGGA ACGATTGAAT TAAGATTGCC AAGTGGAGAA     4860
GTGATAAAAA AGAATTTATC TGATTTCATA GCGTAAGGAA TAGCAGTAGA AAAAGTCTGA     4920
ATCAAAAATG AAGTTCTCTC AAAAGTTAGA AATAAAACTC TGACTTTGGG AGAATTTCAT     4980
TTTATTATTA ATATATAAAA TTTCTTGACA TACAACTTAA AAAGAGGTGG AATATTTACT     5040
AGTTAATT   (SEQ ID NO : 2)                                            5048
```

FIGURE 2

```
ATGAAAATCA ATAAAAAATA TCTAGCTGGG TCAGTAGCTA CACTTGTTTT AAGTGTCTGT       60
GCTTATGAAC TAGGTTTGCA TCAAGCTCAA ACTGTAAAAG AAAATAATCG TGTTTCCTAT      120
ATAGATGGAA AACAAGCGAC GCAAAAAACG GAGAATTTGA CTCCTGATGA GGTTAGCAAG      180
CGTGAAGGAA TCAACGCCGA ACAAATCGTC ATCAAGATTA CGGATCAAGG TTATGTGACC      240
TCTCATGGAG ACCATTATCA TTACTATAAT GGCAAGGTCC CTTATGATGC CATCATCAGT      300
GAAGAGCTCC TCATGAAAGA TCCGAATTAT CAGTTGAAGG ATTCAGACAT TGTCAATGAA      360
ATCAAGGGTG GTTATGTCAT TAAGGTAAAC GGTAAATACT ATGTTTACCT TAAGGATGCA      420
GCTCATGCGG ATAATGTCCG TACAAAAGAA GAAATCAATC GGCAAAAACA AGAACATAGT      480
CAGCATCGTG AAGGAGGGAC TTCAGCAAAC GATGGTGCGG TAGCCTTTGC ACGTTCACAG      540
GGACGCTACA CCACAGATGA TGGTTATATC TTCAATGCAT CTGATATCAT CGAAGATACG      600
GGCGATGCCT ATATCGTTCC TCATGGAGAT CATTACCATT ACATTCCTAA GAATGAGTTA      660
TCAGCTAGCG AGTTGGCTGC TGCAGAAGCC TTCCTATCTG GTCGGGAAAA TCTGTCAAAT      720
TTAAGAACCT ATCGCCGACA AAATAGCGAT AACACTCCAA GAACAAACTG GGTACCTTCT      780
GTAAGCAATC CAGGAACTAC AAATACTAAC ACAAGCAACA ACAGCAACAC TAACAGTCAA      840
GCAAGTCAAA GTAATGACAT TGATAGTCTC TTGAAACAGC TCTACAAACT GCCTTTGAGT      900
CAACGCCATG TAGAATCTGA TGGCCTTATT TTCGACCCAG CGCAAATCAC AAGTCGAACC      960
GCCAGAGGTG TAGCTGTCCC TCATGGTAAC CATTACCACT TTATCCCTTA TGAACAAATG     1020
TCTGAATTGG AAAAACGAAT TGCTCGTATT ATTCCCCTTC GTTATCGTTC AAACCATTGG     1080
GTACCAGATT CAAGACCAGA AGAACCAAGT CCACAACCGA CTCCAGAACC TAGTCCAAGT     1140
CCGCAACCTG CACCAAATCC TCAACCAGCT CCAAGCAATC CAATTGATGA GAAATTGGTC     1200
AAAGAAGCTG TTCGAAAAGT AGGCGATGGT TATGTCTTTG AGGAGAATGG AGTTTCTCGT     1260
TATATCCCAG CCAAGAATCT TTCAGCAGAA ACAGCAGCAG GCATTGATAG CAAACTGGCC     1320
AAGCAGGAAA GTTTATCTCA TAAGCTAGGA GCTAAGAAAA CTGACCTCCC ATCTAGTGAT     1380
CGAGAATTTT ACAATAAGGC TTATGACTTA CTAGCAAGAA TTCACCAAGA TTTACTTGAT     1440
AATAAAGGTC GACAAGTTGA TTTTGAGGCT TTGGATAACC TGTTGGAACG ACTCAAGGAT     1500
GTCTCAAGTG ATAAAGTCAA GTTAGTGGAT GATATTCTTG CCTTCTTAGC TCCGATTCGT     1560
CATCCAGAAC GTTTAGGAAA ACCAAATGCG CAAATTACCT ACACTGATGA TGAGATTCAA     1620
GTAGCCAAGT TGGCAGGCAA GTACACAACA GAAGACGGTT ATATCTTTGA TCCTCGTGAT     1680
ATAACCAGTG ATGAGGGGGA TGCCTATGTA ACTCCACATA TGACCCATAG CCACTGGATT     1740
AAAAAAGATA GTTTGTCTGA AGCTGAGAGA GCGGCAGCCC AGGCTTATGC TAAAGAGAAA     1800
GGTTTGACCC CTCCTTCGAC AGACCATCAG GATTCAGGAA ATACTGAGGC AAAAGGAGCA     1860
GAAGCTATCT ACAACCGCGT GAAAGCAGCT AAGAAGGTGC CACTTGATCG TATGCCTTAC     1920
AATCTTCAAT ATACTGTAGA AGTCAAAAAC GGTAGTTTAA TCATACCTCA TTATGACCAT     1980
TACCATAACA TCAAATTTGA GTGGTTTGAC GAAGGCCTTT ATGAGGCACC TAAGGGGTAT     2040
ACTCTTGAGG ATCTTTTGGC GACTGTCAAG TACTATGTCG AACATCCAAA CGAACGTCCG     2100
CATTCAGATA ATGGTTTTGG TAACGCTAGC GACCATGTTC AAAGAAACAA AAATGGTCAA     2160
```

GCTGATACCA ATCAAACGGA AAAACCAAGC GAGGAGAAAC CTCAGACAGA AAAACCTGAG 2220
GAAGAAACCC CTCGAGAAGA GAAACCACAA AGCGAGAAAC CAGAGTCTCC AAAACCAACA 2280
GAGGAACCAG AAGAAGAATC ACCAGAGGAA TCAGAAGAAC CTCAGGTCGA GACTGAAAAG 2340
GTTGAAGAAA AACTGAGAGA GGCTGAAGAT TTACTTGGAA AAATCCAGGA TCCAATTATC 2400
AAGTCCAATG CCAAAGAGAC TCTCACAGGA TTAAAAAATA ATTTACTATT TGGCACCCAG 2460
GACAACAATA CTATTATGGC AGAAGCTGAA AAACTATTGG CTTTATTAAA GGAGAGTAAG 2520
TAA (SEQ ID NO: 3) 2523
FIGURE 3

CAGAGATCTT AGTGAATCAA ATATACTTAA GAAAAGAGGA AAGAATGAAA ATCAATAAAA 60
AATATCTAGC TGGGTCAGTA GCTACACTTG TTTTAAGTGT CTGTGCTTAT GAACTAGGTT 120
TGCATCAAGC TCAAACTGTA AAAGAAAATA ATCGTGTTTC CTATATAGAT GGAAAACAAG 180
CGACGCAAAA AACGGAGAAT TTGACTCCTG ATGAGGTTAG CAAGCGTGAA GGAATCAACG 240
CCGAACAAAT CGTCATCAAG ATTACGGATC AAGGTTATGT GACCTCTCAT GGAGACCATT 300
ATCATTACTA TAATGGCAAG GTCCCTTATG ATGCCATCAT CAGTGAAGAG CTCCTCATGA 360
AAGATCCGAA TTATCAGTTG AAGGATTCAG ACATTGTCAA TGAAATCAAG GGTGGTTATG 420
TCATTAAGGT AAACGGTAAA TACTATGTTT ACCTTAAGGA TGCAGCTCAT GCGGATAATG 480
TCCGTACAAA AGAAGAAATC AATCGGCAAA AACAAGAACA TAGTCAGCAT CGTGAAGGAG 540
GGACTTCAGC AAACGATGGT GCGGTAGCCT TTGCACGTTC ACAGGGACGC TACACCACAG 600
ATGATGGTTA TATCTTCAAT GCATCTGATA TCATCGAAGA TACGGGCGAT GCCTATATCG 660
TTCCTCATGG AGATCATTAC CATTACATTC CTAAGAATGA GTTATCAGCT AGCGAGTTGG 720
CTGCTGCAGA AGCCTTCCTA TCTGGTCGGG AAAATCTGTC AAATTTAAGA ACCTATCGCC 780
GACAAAATAG CGATAACACT CCAAGAACAA ACTGGGTACC TTCTGTAAGC AATCCAGGAA 840
CTACAAATAC TAACACAAGC AACAACAGCA ACACTAACAG TCAAGCAAGT CAAAGTAATG 900
ACATTGATAG TCTCTTGAAA CAGCTCTACA AACTGCCTTT GAGTCAACGC CATGTAGAAT 960
CTGATGGCCT TATTTTCGAC CCAGCGCAAA TCACAAGTCG AACCGCCAGA GGTGTAGCTG 1020
TCCCTCATGG TAACCATTAC CACTTTATCC CTTATGAACA AATGTCTGAA TTGGAAAAAC 1080
GAATTGCTCG TATTATTCCC CTTCGTTATC GTTCAAACCA TTGGGTACCA GATTCAAGAC 1140
CAGAAGAACC AAGTCCACAA CCGACTCCAG AACCTAGTCC AAGTCCGCAA CCTGCACCAA 1200
ATCCTCAACC AGCTCCAAGC AATCCAATTG ATGAGAAATT GGTCAAAGAA GCTGTTCGAA 1260
AAGTAGGCGA TGGTTATGTC TTTGAGGAGA ATGGAGTTTC TCGTTATATC CCAGCCAAGA 1320
ATCTTTCAGC AGAAACAGCA GCAGGCATTG ATAGCAAACT GGCCAAGCAG GAAAGTTTAT 1380
CTCATAAGCT AGGAGCTAAG AAAACTGACC TCCCATCTAG TGATCGAGAA TTTTACAATA 1440
AGGCTTATGA CTTACTAGCA AGAATTCACC AAGATTTACT TGATAATAAA GGTCGACAAG 1500
TTGATTTTGA GGCTTTGGAT AACCTGTTGG AACGACTCAA GGATGTCTCA AGTGATAAAG 1560
TCAAGTTAGT GGATGATATT CTTGCCTTCT TAGCTCCGAT TCGTCATCCA GAACGTTTAG 1620
GAAAACCAAA TGCGCAAATT ACCTACACTG ATGATGAGAT TCAAGTAGCC AAGTTGGCAG 1680
GCAAGTACAC AACAGAAGAC GGTTATATCT TTGATCCTCG TGATATAACC AGTGATGAGG 1740
GGGATGCCTA TGTAACTCCA CATATGACCC ATAGCCACTG GATTAAAAAA GATAGTTTGT 1800
CTGAAGCTGA GAGAGCGGCA GCCCAGGCTT ATGCTAAAGA GAAAGGTTTG ACCCCTCCTT 1860
CGACAGACCA TCAGGATTCA GGAAATACTG AGGCAAAAGG AGCAGAAGCT ATCTACAACC 1920
GCGTGAAAGC AGCTAAGAAG GTGCCACTTG ATCGTATGCC TTACAATCTT CAATATACTG 1980
TAGAAGTCAA AAACGGTAGT TTAATCATAC CTCATTATGA CCATTACCAT AACATCAAAT 2040
TTGAGTGGTT TGACGAAGGC CTTTATGAGG CACCTAAGGG GTATACTCTT GAGGATCTTT 2100

TGGCGACTGT CAAGTACTAT GTCGAACATC CAAACGAACG TCCGCATTCA GATAATGGTT 2160
TTGGTAACGC TAGCGACCAT GTTCAAAGAA ACAAAAATGG TCAAGCTGAT ACCAATCAAA 2220
CGGAAAAACC AAGCGAGGAG AAACCTCAGA CAGAAAAACC TGAGGAAGAA ACCCCTCGAG 2280
AAGAGAAACC ACAAAGCGAG AAACCAGAGT CTCCAAAACC AACAGAGGAA CCAGAAGAAG 2340
AATCACCAGA GGAATCAGAA GAACCTCAGG TCGAGACTGA AAAGGTTGAA GAAAAACTGA 2400
GAGAGGCTGA AGATTTACTT GGAAAAATCC AGGATCCAAT TATCAAGTCC AATGCCAAAG 2460
AGACTCTCAC AGGATTAAAA AATAATTTAC TATTTGGCAC CCAGGACAAC AATACTATTA 2520
TGGCAGAAGC TGAAAAACTA TTGGCTTTAT TAAAGGAGAG TAAGTAAAGG TAGCAGCATT 2580
TTCTAACTCC TAAAAACAGG ATAGGAGAAC GGGAAAACGA AAAATGAGAG CAGAATGTGA 2640
GTTCTAG (SED ID NO : 4) 2647
FIGURE 4

GGGTCTTAAA ACTCTGAATC CTTTAGAGGC AGACCCACAA AATGACAAGA CCTATTTAGA 60
AAATCTGGAA GAAAATATGA GTGTTCTAGC AGAAGAATTA AAGTGAGGAA AGAATGAAAA 120
TCAATAAAAA ATATCTAGCA GGTTCAGTGG CAGTCCTTGC CCTAAGTGTT TGTTCCTATG 180
AACTTGGTCG TCACCAAGCT GGTCAGGTTA AGAAAGAGTC TAATCGAGTT TCTTATATAG 240
ATGGTGATCA GGCTGGTCAA AAGGCAGAAA ATTTGACACC AGATGAAGTC AGTAAGAGAG 300
AGGGGATCAA CGCCGAACAA ATTGTTATCA AGATTACGGA TCAAGGTTAT GTGACCTCTC 360
ATGGAGACCA TTATCATTAC TATAATGGCA AGGTTCCTTA TGATGCCATC ATCAGTGAAG 420
AACTTCTCAT GAAAGATCCG AATTATCAGT TGAAGGATTC AGACATTGTC AATGAAATCA 480
AGGGTGGCTA TGTGATTAAG GTAGACGGAA AATACTATGT TTACCTTAAA GATGCGGCCC 540
ATGCGGACAA TATTCGGACA AAAGAAGAGA TTAAACGTCA GAAGCAGGAA CACAGTCATA 600
ATCATAACTC AAGAGCAGAT AATGCTGTTG CTGCAGCCAG AGCCCAAGGA CGTTATACAA 660
CGGATGATGG GTATATCTTC AATGCATCTG ATATCATTGA GGACACGGGT GATGCTTATA 720
TCGTTCCTCA CGGCGACCAT TACCATTACA TTCCTAAGAA TGAGTTATCA GCTAGCGAGT 780
TAGCTGCTGC AGAAGCCTAT TGGAATGGGA AGCAGGGATC TCGTCCTTCT TCAAGTTCTA 840
GTTATAATGC AAATCCAGTT CAACCAAGAT TGTCAGAGAA CCACAATCTG ACTGTCACTC 900
CAACTTATCA TCAAAATCAA GGGGAAAACA TTTCAAGCCT TTTACGTGAA TTGTATGCTA 960
AACCCTTATC AGAACGCCAT GTAGAATCTG ATGGCCTTAT TTTCGACCCA GCGCAAATCA 1020
CAAGTCGAAC CGCCAGAGGT GTAGCTGTCC CTCATGGTAA CCATTACCAC TTTATCCCTT 1080
ATGAACAAAT GTCTGAATTG GAAAAACGAA TTGCTCGTAT TATTCCCCTT CGTTATCGTT 1140
CAAACCATTG GGTACCAGAT TCAAGACCAG AACAACCAAG TCCACAATCG ACTCCGGAAC 1200
CTAGTCCAAG TCTGCAACCT GCACCAAATC CTCAACCAGC TCCAAGCAAT CCAATTGATG 1260
AGAAATTGGT CAAAGAAGCT GTTCGAAAAG TAGGCGATGG TTATGTCTTT GAGGAGAATG 1320
GAGTTTCTCG TTATATCCCA GCCAAGGATC TTTCAGCAGA AACAGCAGCA GGCATTGATA 1380
GCAAACTGGC CAAGCAGGAA AGTTTATCTC ATAAGCTAGG AGCTAAGAAA ACTGACCTCC 1440
CATCTAGTGA TCGAGAATTT TACAATAAGG CTTATGACTT ACTAGCAAGA ATTCACCAAG 1500
ATTTACTTGA TAATAAAGGT CGACAAGTTG ATTTTGAGGT TTTGGATAAC CTGTTGGAAC 1560
GACTCAAGGA TGTCTCAAGT GATAAAGTCA AGTTAGTGGA TGATATTCTT GCCTTCTTAG 1620
CTCCGATTCG TCATCCAGAA CGTTTAGGAA AACCAAATGC GCAAATTACC TACACTGATG 1680
ATGAGATTCA AGTAGCCAAG TTGGCAGGCA AGTACACAAC AGAAGACGGT TATATCTTTG 1740
ATCCTCGTGA TATAACCAGT GATGAGGGGG ATGCCTATGT AACTCCACAT ATGACCCATA 1800
GCCACTGGAT TAAAAAAGAT AGTTTGTCTG AAGCTGAGAG AGCGGCAGCC CAGGCTTATG 1860
CTAAAGAGAA AGGTTTGACC CCTCCTTCGA CAGACCACCA GGATTCAGGA AATACTGAGG 1920

```
CAAAAGGAGC AGAAGCTATC TACAACCGCG TGAAAGCAGC TAAGAAGGTG CCACTTGATC     1980
GTATGCCTTA CAATCTTCAA TATACTGTAG AAGTCAAAAA CGGTAGTTTA ATCATACCTC     2040
ATTATGACCA TTACCATAAC ATCAAATTTG AGTGGTTTGA CGAAGGCCTT TATGAGGCAC     2100
CTAAGGGGTA TAGTCTTGAG GATCTTTTGG CGACTGTCAA GTACTATGTC GAACATCCAA     2160
ACGAACGTCC GCATTCAGAT AATGGTTTTG GTAACGCTAG TGACCATGTT CGTAAAAATA     2220
AGGCAGACCA AGATAGTAAA CCTGATGAAG ATAAGGAACA TGATGAAGTA AGTGAGCCAA     2280
CTCACCCTGA ATCTGATGAA AAAGAGAATC ACGCTGGTTT AAATCCTTCA GCAGATAATC     2340
TTTATAAACC AAGCACTGAT ACGGAAGAGA CAGAGGAAGA AGCTGAAGAT ACCACAGATG     2400
AGGCTGAAAT TCCTCAAGTA GAGAATTCTG TTATTAACGC TAAGATAGCA GATGCGGAGG     2460
CCTTGCTAGA AAAAGTAACA GATCCTAGTA TTAGACAAAA TGCTATGGAG ACATTGACTG     2520
GTCTAAAAAG TAGTCTTCTT CTCGGAACGA AAGATAATAA CACTATTTCA GCAGAAGTAG     2580
ATAGTCTCTT GGCTTTGTTA AAAGAAAGTC AACCGGCTCC TATACAGTAG TAAAATGAA      2639
```

(SEQ ID NO : 5)

FIGURE 5

```
MKFSKKYIAA GSAVIVSLSL CAYALNQHRS QENKDNNRVS YVDGSQSSQK       50
SENLTPDQVS QKEGIQAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDALF      100
SEELLMKDPN YQLKDADIVN EVKGGYIIKV DGKYYVYLKD AAHADNVRTK      150
DEINRQKQEH VKDNEKVNSN VAVARSQGRY TTNDGYVFNP ADIIEDTGNA      200
YIVPHGGHYH YIPKSDLSAS ELAAAKAHLA GKNMQPSQLS YSSTASDNNT      250
QSVAKGSTSK PANKSENLQS LLKELYDSPS AQRYSESDGL VFDPAKIISR      300
TPNGVAIPHG DHYHFIPYSK LSALEEKIAR MVPISGTGST VSTNAKPNEV      350
VSSLGSLSSN PSSLTTSKEL SSASDGYIFN PKDIVEETAT AYIVRHGDHF      400
HYIPKSNQIG QPTLPNNSLA TPSPSLPINP GTSHEKHEED GYGFDANRII      450
AEDESGFVMS HGDHNHYFFK KDLTEEQIKA AQKHLEEVKT SHNGLDSLSS      500
HEQDYPGNAK EMKDLDKKIE EKIAGIMKQY GVKRESIVVN KEKNAIIYPH      550
GDHHHADPID EHKPVGIGHS HSNYELFKPE EGVAKKEGNK VYTGEELTNV      600
VNLLKNSTFN NQNFTLANGQ KRVSFSFPPE LEKKLGINML VKLITPDGKV      650
LEKVSGKVFG EGVGNIANFE LDQPYLPGQT FKYTIASKDY PEVSYDGTFT      700
VPTSLAYKMA SQTIFYPFHA GDTYLRVNPQ FAVPKGTDAL VRVFDEFHGN      750
AYLENNYKVG EIKLPIPKLN QGTTRTAGNK IPVTFMANAY LDNQSTYIVE      800
VPILEKENQT DKPSILPQFK RNKAQENSKL DEKVEEPKTS EKVEKEKLSE      850
TGNSTSNSTL EEVPTVDPVQ EKVAKFAESY GMKLENVLFN MDGTIELYLP      900
SGEVIKKNMA DFTGEAPQGN GENKPSENGK VSTGTVENQP TENKPADSLP      950
EAPNEKPVKP ENSTDNGMLN PEGNVGSDPM LDPALEEAPA VDPVQEKLEK     1000
FTASYGLGLD SVIFNMDGTI ELRLPSGEVI KKNLSDFIA (SEQ ID NO: 6)  1039
```

FIGURE 6

```
MKINKKYLAG SVATLVLSVC AYELGLHQAQ TVKENNRVSY IDGKQATQKT       50
ENLTPDEVSK REGINAEQIV IKITDQGYVT SHGDHYHYYN GKVPYDAIIS      100
EELLMKDPNY QLKDSDIVNE IKGGYVIKVN GKYYVYLKDA AHADNVRTKE      150
EINRQKQEHS QHREGGTSAN DGAVAFARSQ GRYTTDDGYI FNASDIIEDT      200
GDAYIVPHGD HYHYIPKNEL SASELAAAEA FLSGRENLSN LRTYRRQNSD      250
```

```
NTPRTNWVPS VSNPGTTNTN TSNNSNTNSQ ASQSNDIDSL LKQLYKLPLS      300
QRHVESDGLI FDPAQITSRT ARGVAVPHGN HYHFIPYEQM SELEKRIARI      350
IPLRYRSNHW VPDSRPEEPS PQPTPEPSPS PQPAPNPQPA PSNPIDEKLV      400
KEAVRKVGDG YVFEENGVSR YIPAKNLSAE TAAGIDSKLA KQESLSHKLG      450
AKKTDLPSSD REFYNKAYDL LARIHQDLLD NKGRQVDFEA LDNLLERLKD      500
VSSDKVKLVD DILAFLAPIR HPERLGKPNA QITYTDDEIQ VAKLAGKYTT      550
EDGYIFDPRD ITSDEGDAYV TPHMTHSHWI KKDSLSEAER AAAQAYAKEK      600
GLTPPSTDHQ DSGNTEAKGA EAIYNRVKAA KKVPLDRMPY NLQYTVEVKN      650
GSLIIPHYDH YHNIKPEWFD EGLYEAPKGY TLEDLLATVK YYVEHPNERP      700
HSDNGFGNAS DHVQRNKNGQ ADTNQTEKPS EEKPQTEKPE EETPREEKPQ      750
SEKPESPKPT EEPEEESPEE SEEPQVETEK VEEKLREAED LLGKIQDPII      800
KSNAKETLTG LKNNLLFGTQ DNNTIMAEAE KLLALLKESK (SEQ ID NO: 7)  840
```

FIGURE 7

```
MKINKKYLAG SVAVLALSVC SYELGRHQAG QVKKESNRVS YIDGDQAGQK       50
AENLTPDEVS KREGINAEQI VIKITDQGYV TSHGDHYHYY NGKVPYDAII      100
SEELLMKDPN YQLKDSDIVN EIKGGYVIKV DGKYYVYLKD AAHADNIRTK      150
EEIKRQKQEH SHNHNSRADN AVAAARAQGR YTTDDGYIFN ASDIIEDTGD      200
AYIVPHGDHY HYIPKNELSA SELAAAEAYW NGKQGSRPSS SSSYNANPVQ      250
PRLSENHNLT VTPTYHQNQG ENISSLLREL YAKPLSERHV ESDGLIFDPA      300
QITSRTARGV AVPHGNHYHF IPYEQMSELE KRIARIIPLR YRSNHWVPDS      350
RPEQPSPQST PEPSPSLQPA PNPQPAPSNP IDEKLVKEAV RKVGDGYVFE      400
ENGVSRYIPA KDLSAETAAG IDSKLAKQES LSHKLGAKKT DLPSSDREFY      450
NKAYDLLARI HQDLLDNKGR QVDFEVLDNL LERLKDVSSD KVKLVDDILA      500
FLAPIRHPER LGKPNAQITY TDDEIQVAKL AGKYTTEDGY IFDPRDITSD      550
EGDAYVTPHM THSHWIKKDS LSEAERAAAQ AYAKEKGLTP PSTDHQDSGN      600
TEAKGAEAIY NRVKAAKKVP LDRMPYNLQY TVEVKNGSLI IPHYDHYHNI      650
KFEWFDEGLY EAPKGYSLED LLATVKYYVE HPNERPHSDN GFGNASDHVR      700
KNKADQDSKP DEDKEHDEVS EPTHPESDEK ENHAGLNPSA DNLYKPSTDT      750
EETEEEAEDT TDEAEIPQVE NSVINAKIAD AEALLEKVTD PSIRQNAMET      800
LTGLKSSLLL GTKDNNTISA EVDSLLALLK ESQPAPIQ                   838
```

(SEQ ID NO : 8)

FIGURE 8

```
TGTGCCTATG CACTAAACCA GCATCGTTCG CAGGAAAATA AGGACAATAA TCGTGTCTCT       60
TATGTGGATG GCAGCCAGTC AAGTCAGAAA AGTGAAAACT TGACACCAGA CCAGGTTAGC      120
CAGAAAGAAG GAATTCAGGC TGAGCAAATT GTAATCAAAA TTACAGATCA GGGCTATGTA      180
ACGTCACACG GTGATCACTA TCATTACTAT AATGGGAAAG TTCCTTATGA TGCCCTCTTT      240
AGTGAAGAAC TCTTGATGAA GGATCCAAAC TATCAACTTA AAGACGCTGA TATTGTCAAT      300
GAAGTCAAGG GTGGTTATAT CATCAAGGTC GATGGAAAAT ATTATGTCTA CCTGAAAGAT      360
GCAGCTCATG CTGATAATGT TCGAACTAAA GATGAAATCA ATCGTCAAAA ACAAGAACAT      420
GTCAAAGATA ATGAGAAGGT TAACTCTAAT GTTGCTGTAG CAAGGTCTCA GGGACGATAT      480
ACGACAAATG ATGGTTATGT CTTTAATCCA GCTGATATTA TCGAAGATAC GGGTAATGCT      540
```

```
TATATCGTTC CTCATGGAGG TCACTATCAC TACATTCCCA AAAGCGATTT ATCTGCTAGT      600
GAATTAGCAG CAGCTAAAGC ACATCTGGCT GGAAAAAATA TGCAACCGAG TCAGTTAAGC      660
TATTCTTCAA CACCTTCTCC ATCTCTTCCA ATCAATCCAG GAACTTCACA TGAGAAACAT      720
GAAGAAGATG GATACGGATT TGATGCTAAT CGTATTATCG CTGAAGATGA ATCAGGTTTT      780
GTCATGAGTC ACGGAGACCA CAATCATTAT TTCTTCAAGA AGGACTTGAC AGAAGAGCAA      840
ATTAAGGCTG CGCAAAAACA TTTAGAGGAA GTTAAAACTA GTCATAATGG ATTAGATTCT      900
TTGTCATCTC ATGAACAGGA TTATCCAAGT AATGCCAAAG AAATGAAAGA TTTAGATAAA      960
AAAATCGAAG AAAAAATTGC TGGCATTATG AAACAATATG GTGTCAAACG TGAAAGTATT     1020
GTCGTGAATA AAGAAAAAAA TGCGATTATT TATCCGCATG GAGATCACCA TCATGCAGAT     1080
CCGATTGATG AACATAAACC GGTTGGAATT GGTCATTCTC ACAGTAACTA TGAACTGTTT     1140
AAACCCGAAG AAGGAGTTGC TAAAAAAGAA GGGAATAAAG TTTATACTGG AGAAGAATTA     1200
ACGAATGTTG TTAATTTGTT AAAAAATAGT ACGTTTAATA ATCAAAACTT TACTCTAGCC     1260
AATGGTCAAA AACGCGTTTC TTTTAGTTTT CCGCCTGAAT TGGAGAAAAA ATTAGGTATC     1320
AATATGCTAG TAAAATTAAT AACACCAGAT GGAAAAGTAT TGGAGAAAGT ATCTGGTAAA     1380
GTATTTGGAG AAGGAGTAGG GAATATTGCA AACTTTGAAT TAGATCAACC TTATTTACCA     1440
GGACAAACAT TTAAGTATAC TATCGCTTCA AAAGATTATC CAGAAGTAAG TTATGATGGT     1500
ACATTTACAG TTCCAACCTC TTTAGCTTAC AAAATGGCCA GTCAAACGAT TTTCTATCCT     1560
TTCCATGCAG GGGATACTTA TTTAAGAGTG AACCCTCAAT TTGCAGTGCC TAAAGGAACT     1620
GATGCTTTAG TCAGAGTGTT TGATGAATTT CATGGAAATG CTTATTTAGA AAATAACTAT     1680
AAAGTTGGTG AAATCAAATT ACCGATTCCG AAATTAAACC AAGGAACAAC CAGAACGGCC     1740
GGAAATAAAA TTCCTGTAAC CTTCATGGCA AATGCTTATT TGGACAATCA ATCGACTTAT     1800
ATTGTGGAAG TACCTATCTT GGAAAAAGAA AATCAAACTG ATAAACCAAG TATTCTACCA     1860
CAATTTAAAA GGAATAAAGC ACAAGAAAAC TCAAAACTTG ATGAAAAGGT AGAAGAACCA     1920
AAGACTAGTG AGAAGGTAGA AAAAGAAAAA CTTTCTGAAA CTGGGAATAG TACTAGTAAT     1980
TCAACGTTAG AAGAAGTTCC TACAGTGGAT CCTGTACAAG AAAAAGTAGC AAAATTTGCT     2040
GAAAGTTATG GGATGAAGCT AGAAAATGTC TTGTTTAATA TGGACGGAAC AATTGAATTA     2100
TATTTACCAT CGGGAGAAGT CATTAAAAAG AATATGGCAG ATTTTACAGG AGAAGCACCT     2160
CAAGGAAATG GTGAAAATAA ACCATCTGAA AATGGAAAAG TATCTACTGG AACAGTTGAG     2220
AACCAACCAA CAGAAAATAA ACCAGCAGAT TCTTTACCAG AGGCACCAAA CGAAAAACCT     2280
GTAAAACCAG AAAACTCAAC GGATAATGGA ATGTTGAATC CAGAAGGGAA TGTGGGGAGT     2340
GACCCTATGT TAGATTCAGC ATTAGAGGAA GCTCCAGCAG TAGATCCTGT ACAAGAAAAA     2400
TTAGAAAAAT TTACAGCTAG TTACGGATTA GGCTTAGATA GTGTTATATT CAATATGGAT     2460
GGAACGATTG AATTAAGATT GCCAAGTGGA GAAGTGATAA AAAAGAATTT ATTGATCTCA     2520
TAGCGTAA   (SEQ ID NO : 9)                                            2528
```

FIGURE 9

```
CAYALNQHRS QENKDNNRVS YVDGSQSSQK SENLTPDQVS QKEGIQAEQI       50
VIKITDQGYV TSHGDHYHYY NGKVPYDALF SEELLMKDPN YQLKDADIVN      100
EVKGGYIIKV DGKYYVYLKD AAHADNVRTK DEINRQKQEH VKDNEKVNSN      150
VAVARSQGRY TTNDGYVFNP ADIIEDTGNA YIVPHGGHYH YIPKSDLSAS      200
ELAAAKAHLA GKNMQPSQLS YSSTPSPSLP INPGTSHEKH EEDGYGFDAN      250
RIIAEDESGF VMSHGDHNHY FFKKDLTEEQ IKAAQKHLEE VKTSHNGLDS      300
LSSHEQDYPS NAKEMKDLDK KIEEKIAGIM KQYGVKRESI VVNKEKNAII      350
YPHGDHHHAD PIDEHKPVGI GHSHSNYELF KPEEGVAKKE GNKVYTGEEL      400
```

TNVVNLLKNS TFNNQNFTLA NGQKRVSFSF PPELEKKLGI NMLVKLITPD 450
GKVLEKVSGK VFGEGVGNIA NFELDQPYLP GQTFKYTIAS KDYPEVSYDG 500
TFTVPTSLAY KMASQTIFYP FHAGDTYLRV NPQFAVPKGT DALVRVFDEF 550
HGNAYLENNY KVGEIKLPIP KLNQGTTRTA GNKIPVTFMA NAYLDNQSTY 600
IVEVPILEKE NQTDKPSILP QFKRNKAQEN SKLDEKVEEP KTSEKVEKEK 650
LSETGNSTSN STLEEVPTVD PVQEKVAKFA ESYGMKLENV LFNMDGTIEL 700
YLPSGEVIKK NMADFTGEAP QGNGENKPSE NGKVSTGTVE NQPTENKPAD 750
SLPEAPNEKP VKPENSTDNG MLNPEGNVGS DPMLDSALEE APAVDPVQEK 800
LEKFTASYGL GLDSVIFNMD GTIELRLPSG EVIKKNLLIS 840
(SEQ ID NO : 10)
FIGURE 10

DQGYVTSHGD HYHYYNGKVP YDALFSEELL MKDPNYQLKD ADIVNEVKGG YIIKVDGKYY VYLKDAAHAD NVRTKDEINR QKQEHVKDNE KVNS
(SEQ ID NO: 11)
FIGURE 11

GIQAEQIVIK ITDQGYVTSH GDHYHYYNGK VPYDALFSEE LL
(SEQ ID NO: 12)
FIGURE 12

TAYIVRHGDH FHYIPKSNQI GQPTLPNNSL ATPSPSLPI
(SEQ ID NO: 13)
FIGURE 13

TSNSTLEEVP TVDPVQEKVA KFAESYGMKL ENVLFN
(SEQ ID NO: 14)
FIGURE 14

MDGTIELRLP SGEVIKKNLS DFIA
(SEQ ID NO: 15)
FIGURE 15

YGLGLDSVIF NMDGTIELRL PSGEVIKKNL SDFIA
(SEQ ID NO: 16)
FIGURE 16

PALEEAPAVD PVQEKLEKFT ASYGLGLDSV IFNMDGTIEL RLPSGEVIKK NLSDFIA
(SEQ ID NO: 17)
FIGURE 17

KVEEPKTSEK VEKEKLSETG NSTSNSTLEE VPTVDPVQEK
(SEQ ID NO: 18)
FIGURE 18

MKDLDKKIEE KIAGIMKQYG VKRESIVVNK EKNAIIYPHG DHHHADPIDE HKPVGIGHSH
SNYELFKPEE GVAKKEGN
(SEQ ID NO: 19)
FIGURE19

AIIYPHGDHH HADPIDEHKP VGIGHSHSNY ELFKPEEGVA KKEGNKVYTG E
(SEQ ID NO: 20)
FIGURE 20

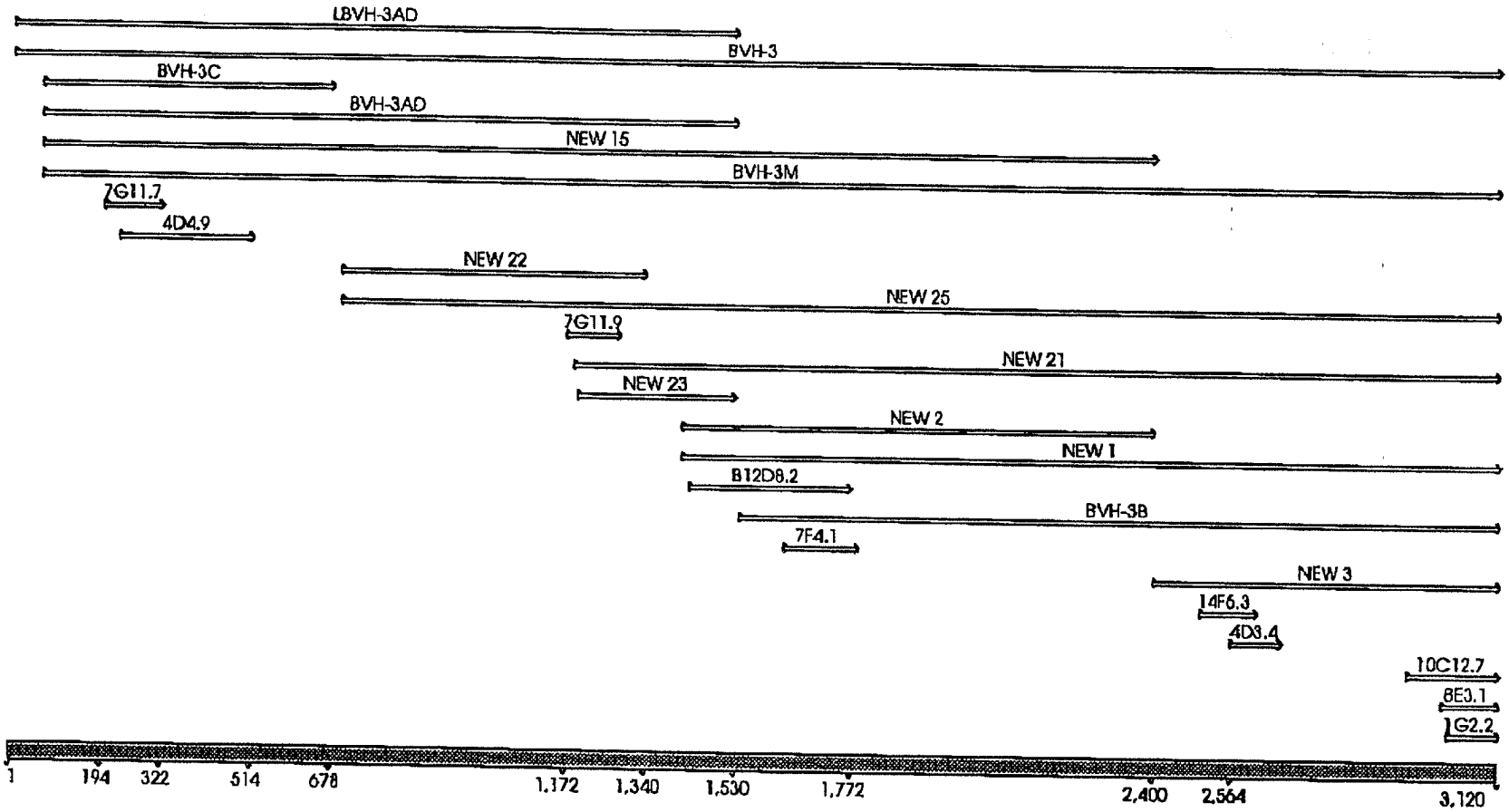
IQVAKLAGKY TTEDGYIFDP RDITSDEGD
(SEQ ID NO: 21)
FIGURE 21

DHQDSGNTEA KGAEAIYNRV KAAKKVPLDR MPYNLQYTVE VKNGSLIIPH YDHYHNIKFE
WFDEGLYEAP KGYSLEDLLA TVKYYV
(SEQ ID NO: 22)
FIGURE 22

GLYEAPKGYS LEDLLATVKY YVEHPNERPH SDNGFGNASD H
(SEQ ID NO: 23)
FIGURE 23

GLYEAPKGYSLEDLLATVKYYV
(SEQ ID NO: 163)
Figure 24

SUBSTITUTE SHEET (RULE 26)



5 FIGURE 25

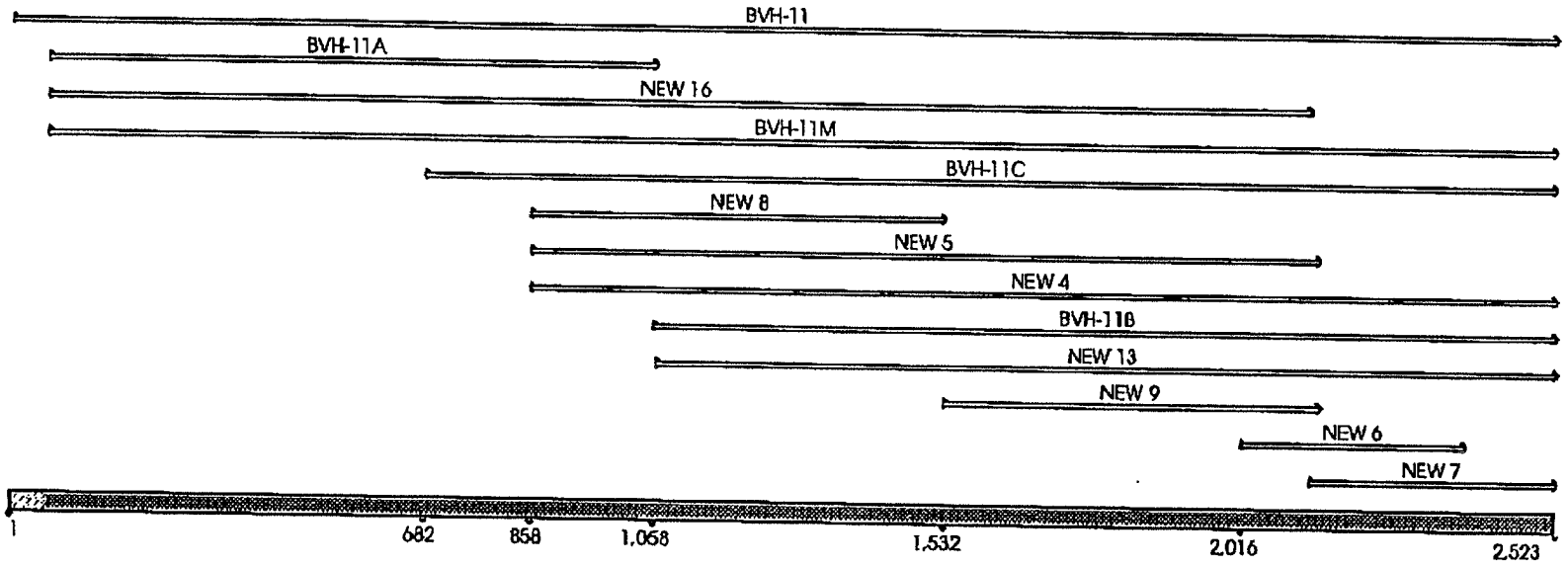

BVH-11
BVH-11A
NEW 16
BVH-11M
BVH-11C
NEW 8
NEW 5
NEW 4
BVH-11B
NEW 13
NEW 9
NEW 6
NEW 7
1
682
858
1,058
1,532
2,016
2,523

FIGURE 26

PCT/US98/03810

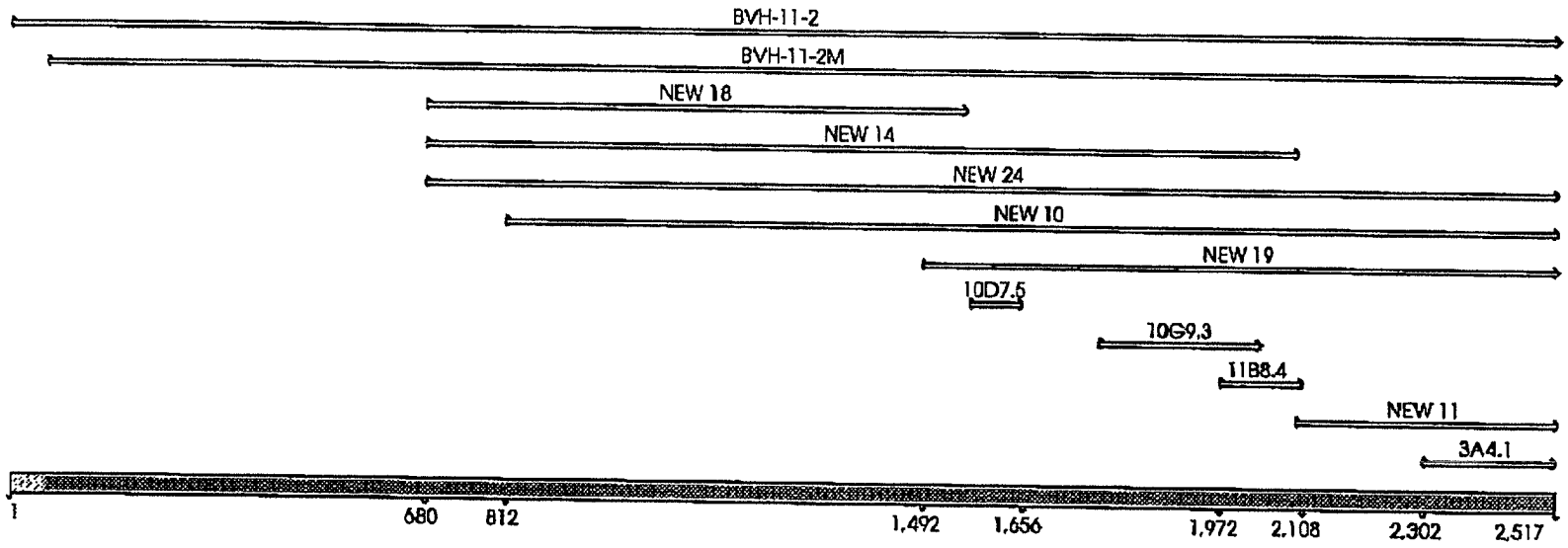


FIGURE 27

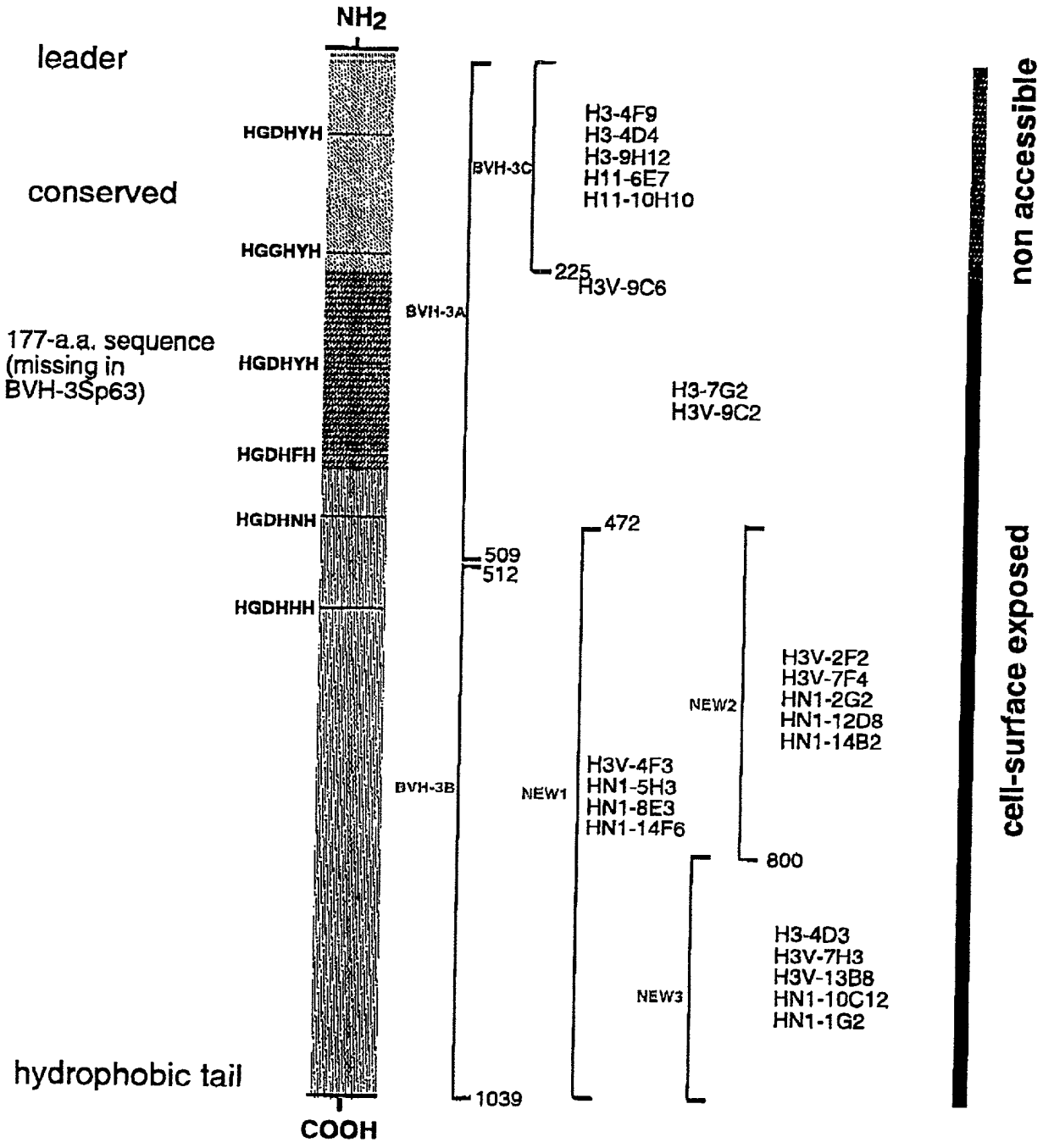
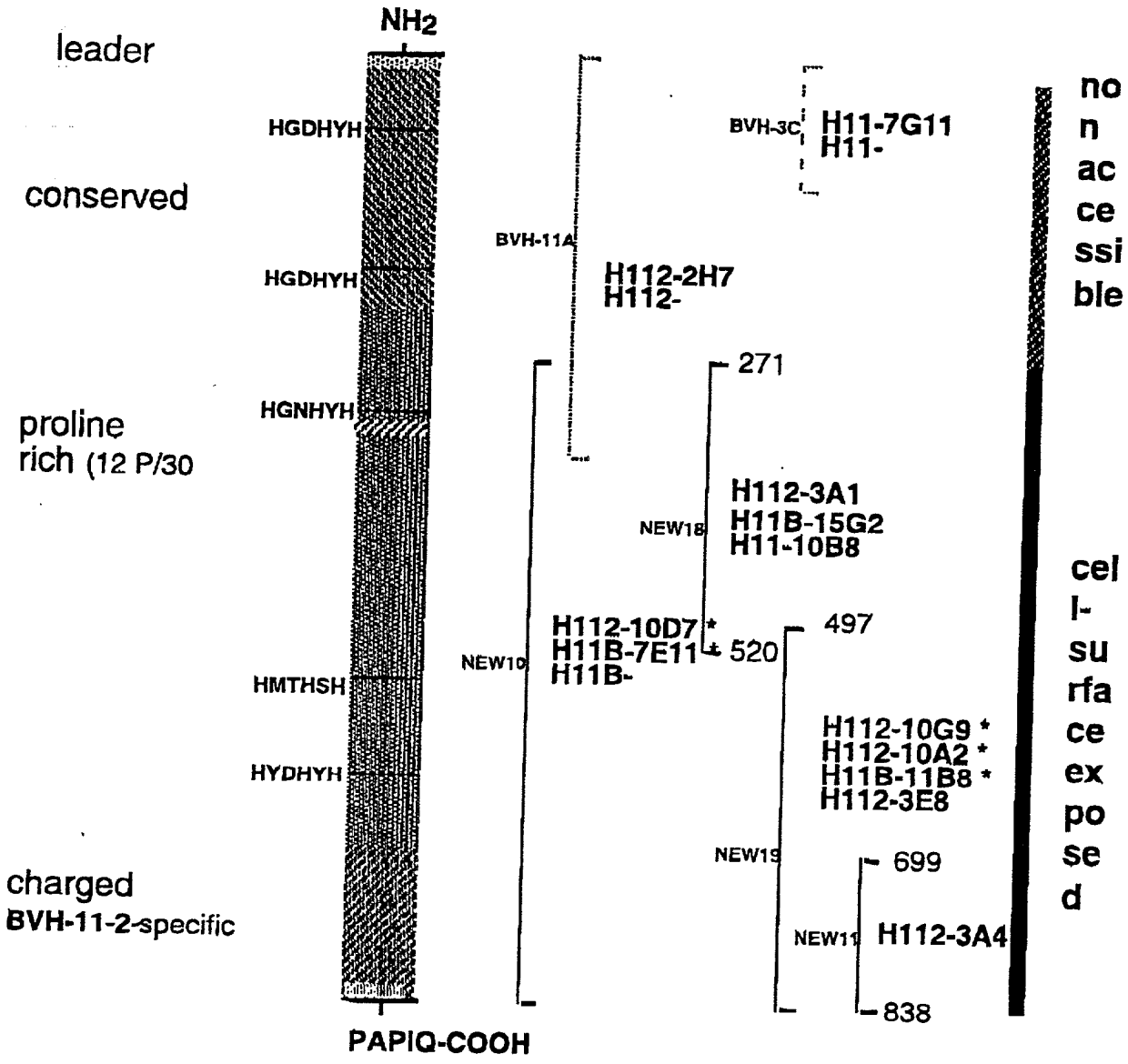

Epitope Localization on BVH-3 Protein
NH2
leader
conserved
177-a.a. sequence
(missing in
BVH-3Sp63)
hydrophobic tail
COOH
HGDHYH
HGGHYH
HGDHYH
HGDHFH
HGDHNH
HGDHHH
BVH-3C
BVH-3A
BVH-3B
H3-4F9
H3-4D4
H3-9H12
H11-6E7
H11-10H10
225
H3V-9C6
H3-7G2
H3V-9C2
509
512
472
NEW1
H3V-4F3
HN1-5H3
HN1-8E3
HN1-14F6
NEW2
H3V-2F2
H3V-7F4
HN1-2G2
HN1-12D8
HN1-14B2
800
NEW3
H3-4D3
H3V-7H3
H3V-13B8
HN1-10C12
HN1-1G2
1039
non accessible
cell-surface exposed

FIGURE 28

Epitope Localization on BVH-11-2 Protein



* Surface-exposed and protection-conferring Mabs

FIGURE 29

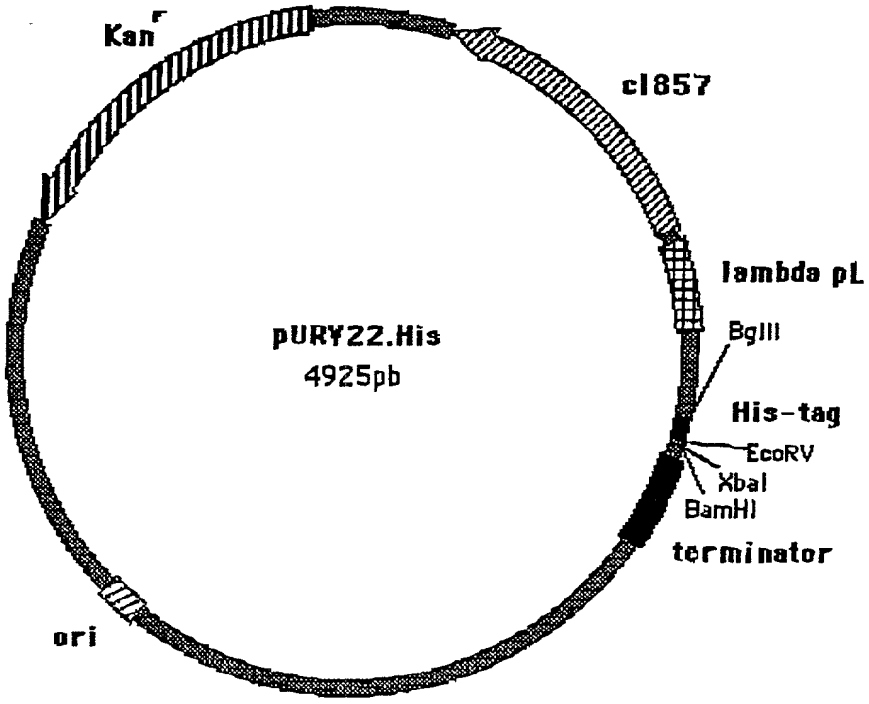

Kan
cI857
lambda pL
BglII
His-tag
EcoRV
XbaI
BamHI
terminator
ori
pURY22.His
4925pb

FIGURE 30

```
BVH-3M      1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV  60
BVH3-63     1 CAYALNQHRSQENKDNNRVSYVDGSQSSQKSENLTPDQVSQKEGIQAEQIVIKITDQGYV  60
              ************************************************************

BVH-3M     61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 120
BVH3-63    61 TSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIVNEVKGGYIIKVDGKYYVYLKD 120
              ************************************************************

BVH-3M    121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 180
BVH3-63   121 AAHADNVRTKDEINRQKQEHVKDNEKVNSNVAVARSQGRYTTNDGYVFNPADIIEDTGNA 180
              ************************************************************

BVH-3M    181 YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSSTASDNNTQSVAKGSTSK 240
BVH3-63   181 YIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQPSQLSYSS---------------- 223
              ****************************************

BVH-3M    241 PANKSENLQSLLKELYDSPSAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSK 300
BVH3-63   224 ------------------------------------------------------------ 223

BVH-3M    301 LSALEEKIARMVPISGTGSTVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFN 360
BVH3-63   224 ------------------------------------------------------------ 223

BVH-3M    361 PKDIVEETATAYIVRHGDHFHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEED 420
BVH3-63   224 ----------------------------------TPSPSLPINPGTSHEKHEED 243
                                                 ********************

BVH-3M    421 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 480
BVH3-63   244 GYGFDANRIIAEDESGFVMSHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSS 303
              ************************************************************

BVH-3M    481 HEQDYPGNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 540
BVH3-63   304 HEQDYPSNAKEMKDLDKKIEEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPID 363
              ****** ***************************************************

BVH-3M    541 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 600
BVH3-63   364 EHKPVGIGHSHSNYELFKPEEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQ 423
              ************************************************************

BVH-3M    601 KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 660
BVH3-63   424 KRVSFSFPPELEKKLGINMLVKLITPDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQT 483
```

```
                ************************************************************

BVH-3M    661 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  720
BVH3-63   484 FKYTIASKDYPEVSYDGTFTVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDAL  543
                ************************************************************

BVH-3M    721 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  780
BVH3-63   544 VRVFDEFHGNAYLENNYKVGEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVE  603
                ************************************************************

BVH-3M    781 VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  840
BVH3-63   604 VPILEKENQTDKPSILPQFKRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTL  663
                ************************************************************

BVH-3M    841 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  900
BVH3-63   664 EEVPTVDPVQEKVAKFAESYGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGN  723
                ************************************************************

BVH-3M    901 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  960
BVH3-63   724 GENKPSENGKVSTGTVENQPTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPM  783
                ************************************************************

BVH-3M    961 LDPALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDFIA 1019
BVH3-63   784 LDSALEEAPAVDPVQEKLEKFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLLIS     840
                ** ***************************************************
```

FIGURE 31

```
BVH-3       1 MKFSKKYIAAGSAVIVSLSLCAYALNQHRSQENK-DNNRVSYVDGSQSSQKSENLTPDQV   59
BVH-11      1 MKINKKYLAG-SVATLVLSVCAYELGLHQAQTVK-ENNRVSYIDGKQATQKTENLTPDEV   58
BVH-11-2    1 MKINKKYLAG-SVAVLALSVCSYELGRHQAGQVKKESNRVSYIDGDQAGQKAENLTPDEV   59
                **  ***.*  *   . **.*.* *. *..   * . *****.** *. **.******.*

BVH-3      60 SQKEGIQAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDALFSEELLMKDPNYQLKDADIV  119
BVH-11     59 SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV  118
BVH-11-2   60 SKREGINAEQIVIKITDQGYVTSHGDHYHYYNGKVPYDAIISEELLMKDPNYQLKDSDIV  119
                *..***.**********************************. ***************.***

BVH-3     120 NEVKGGYIIKVDGKYYVYLKDAAHADNVRTKDEINRQKQEHVKDNE----KVNSNVAVAR  175
BVH-11    119 NEIKGGYVIKVNGKYYVYLKDAAHADNVRTKEEINRQKQEHSQHREGGTSANDGAVAFAR  178
BVH-11-2  120 NEIKGGYVIKVDGKYYVYLKDAAHADNIRTKEEIKRQKQEHSHNHN---SRADNAVAAAR  176
                **.****.*** ***************.***.** ******  .           ** **
```

```
BVH-3      176 SQGRYTTNDGYVFNPADIIEDTGNAYIVPHGGHYHYIPKSDLSASELAAAKAHLAGKNMQ  235
BVH-11     179 SQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAFLSGRENL  238
BVH-11-2   177 AQGRYTTDDGYIFNASDIIEDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQ--  234
               .****** ***.** .******* ******* ******* .********* *   *.

BVH-3      236 PSQLSYSSTASDNNTQSVAKGSTSKP------A-N---------KSENLQSLLKELYDSP  279
BVH-11     239 SNLRTYRRQNSDNTPRTNWVPSVSNPGTTNTNTSNNSNTNSQASQSNDIDSLLKQLYKLP  298
BVH-11-2   235 -GSRPSSSSSYNANPVQPRLSENHNLTVTPTYHQN---------QGENISSLLRELYAKP  284
                              .                *         .   . ***..**  *

BVH-3      280 SAQRYSESDGLVFDPAKIISRTPNGVAIPHGDHYHFIPYSKLSALEEKIARMVPISGTGS  339
BVH-11     299 LSQRHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSN  358
BVH-11-2   285 LSERHVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARIIPLRYRSN  344
                 ..*  *****.****.* ***  ***.*** ******* ..* ** .***..*.

BVH-3      340 TVSTNAKPNEVVSSLGSLSSNPSSLTTSKELSSASDGYIFNPKDIVEETATAYIVRHGDH  399
BVH-11     359 HWVPDSRP-EEPSPQPTPEPSPS-PQPAPNPQPAPS----NP--IDEKLVKEAVRKVGDG  410
BVH-11-2   345 HWVPDSRP-EQPSPQSTPEPSPS-LQPAPNPQPAPS----NP--IDEKLVKEAVRKVGDG  396
                    ..* *  *   .    **    .     *      **  * *      . . **

BVH-3      400 FHYIPKSNQIGQPTLPNNSLATPSPSLPINPGTSHEKHEEDGYGFDANRIIAEDESGFVM  459
BVH-11     411 YVFEE------------NGVSRYIP--------AKNLSAETAAGIDSKLAKQESLS----  446
BVH-11-2   397 YVFEE------------NGVSRYIP--------AKDLSAETAAGIDSKLAKQESLS----  432
                  . .              * ..  *        ..    *   * *.     *  *

BVH-3      460 SHGDHNHYFFKKDLTEEQIKAAQKHLEEVKTSHNGLDSLSSHEQDYPGNAKEMKDLDKKI  519
BVH-11     447 ----HKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALDNLLERLKDVS  502
BVH-11-2   433 ----HKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEVLDNLLERLKDVS  488
                    *     * **        .*  . .   *. *     .. *.      .  *

BVH-3      520 EEKIAGIMKQYGVKRESIVVNKEKNAIIYPHGDHHHADPIDEHKPVGIGHSHSNYELFKP  579
BVH-11     503 SDKVKLVDDILAFLAP--IRHPER--LGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP  558
BVH-11-2   489 SDKVKLVDDILAFLAP--IRHPER--LGKPNAQITYTDDEIQVAKLAGKYTTEDGYIFDP  544
                  .*.  .           . . *.  .  *.     .*   .   .    .    .* *

BVH-3      580 EEGVAKKEGNKVYTGEELTNVVNLLKNSTFNNQNFTLANGQKRVSFSFPPELEKKLGINM  639
BVH-11     559 RD-ITSDEGD-AYVTPHMTHSHWIKKDS-LSEAERAAAQAYAKEKGLTPPSTDHQD----  611
BVH-11-2   545 RD-ITSDEGD-AYVTPHMTHSHWIKKDS-LSEAERAAAQAYAKEKGLTPPSTDHQD----  597
                  . ..  **   *    .*.   . * *       . *.   .      **  ...

BVH-3      640 LVKLITFDGKVLEKVSGKVFGEGVGNIANFELDQPYLPGQTFKYTIASKDYPEVSYDGTF  699
BVH-11     612 ---------------SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVEVKNGSL  653
BVH-11-2   598 ---------------SGNTEAKGAEAIYNRVKAAKKVPLDRMPYNLQ---YTVEVKNGSL  639
                             **     *   * *       .*     *..    *      *.
```

```
BVH-3      700 TVPTSLAYKMASQTIFYPFHAGDTYLRVNPQFAVPKGTDALVRVFDEFHGNAYLENNYKV  759
BVH-11     654 IIP---HYDHYHNIKFEWFDEG--------LYEAPKG--------------YTLEDLLAT  688
BVH-11-2   640 IIP---HYDHYHNIKFEWFDEG--------LYEAPKG--------------YSLEDLLAT  674
                    .*    *    .  *  *  *         .  ***                **

BVH-3      760 GEIKLPIPKLNQGTTRTAGNKIPVTFMANAYLDNQSTYIVEVPILEKENQTDKPSILPQF  819
BVH-11     689 VKYYVEHPNERPHSDNGFGN--------ASDHVQRN-----------KNGQADTN------  724
BVH-11-2   675 VKYYVEHPNERPHSDNGFGN-------ASDHVRKN-----------KADQDSKP------  710
                     .  *     .    **       *   .  .           *  *

BVH-3      820 KRNKAQENSKLDEKVEEPKTSEKVEKEKLSETGNSTSNSTLEEVPTVDPVQEKVAKFAES  879
BVH-11     725 ------QTEKPSEEKPQTEKPE---EE-------------------------------  742
BVH-11-2   711 ------DEDKEHDEVSEPTHPESDEKE-----------------------------  731
                        *  .  .   *    *

BVH-3      880 YGMKLENVLFNMDGTIELYLPSGEVIKKNMADFTGEAPQGNGENKPSENGKVSTGTVENQ  939
BVH-11     743 ------------------------------TPREEKPQSE---KPES----------PK  758
BVH-11-2   732 ------------------------------NHAGLNPSADNLYKPSTD---------TE  751
                                                    *      **

BVH-3      940 PTENKPADSLPEAPNEKPVKPENSTDNGMLNPEGNVGSDPMLDPALEEAPAVDPVQEKLE  999
BVH-11     759 PTEEPEEESPEES--EEPQVETEKVEEKLREAEDLLGK--IQDPIIKSN-----AKETLT  809
BVH-11-2   752 ETEEEAEDTTDEA--EIPQVENSVINAKIADAEALLEK--VTDPSIRQN-----AMETLT  802
                    **    ..  *.  * *          .   *  .    . ** .          * *

BVH-3     1000 KFTASYGLGLDSVIFNMDGTIELRLPSGEVIKKNLSDFIA  1039
BVH-11     810 GLKNNLLFGTQ-----DNNTIMAEAEKLLALLKESK      840
BVH-11-2   803 GLKSSLLLGTK-----DNNTISAEVDSLLALLKESQPAPIQ  838
                       *         .**         . *
```

FIGURE 32

```
  1 ATGCAAATTA CCTACACTGA TGATGAGATT CAGGTAGCCA AGTTGGCAGG CAAGTACACA
 61 ACAGAAGACG GTTATATCTT TGATACTAGT TGGATTAAAA AAGATAGTTT GTCTGAAGCT
121 GAGAGAGCGG CAGCCCAGGC TTATGCTAAA GAGAAAGGTT TGACCCCTCC TTCGACAGAC
181 CACCAGGATT CAGGAAATAC TGAGGCAAAA GGAGCAGAAG CTATCTACAA CCGCGTGAAA
241 GCAGCTAAGA AGGTGCCACT TGATCGTATG CCTTACAATC TTCAGTATAC TGTAGAAGTC
301 AAAAACGGTA GTTTAATCAT ACCTCATTAT GACCATTACC ATAACATCAA ATTTGAGTGG
361 TTTGACGAAG GCCTTTATGA GGCACCTAAG GGGTATAGTC TTGAGGATCT TTTGGCGACT
421 GTCAAGTACT ATGTCGAACC GCGGAACGCT AGTGACCATG TTCGTAAAAA TAAGGCAGAC
481 CAAGATAGTA AACCTGATGA AGATAAGGAA CATGATGAAG TAAGTGAGCC AACTCACCCT
541 GAATCTGATG AAAAAGAGAA TCACGCTGGT TTAAATCCTT CAGCAGATAA TCTTTATAAA
601 CCAAGCACTG ATACGGAAGA GACAGAGGAA GAAGCTGAAG ATACCACAGA TGAGGCTGAA
661 ATTCCTGGTA CCCCTAGTAT TAGACAAAAT GCTATGGAGA CATTGACTGG TCTAAAAAGT
721 AGTCTTCTTC TCGGAACGAA AGATAATAAC ACTATTTCAG CAGAAGTAGA TAGTCTCTTG
781 GCTTTGTTAA AAGAAAGTCA ACCGGCTCCT ATACAGTAG  (SEQ ID NO: 257)
```

FIGURE 33

```
  1 MQITYTDDEI QVAKLAGKYT TEDGYIFDTS WIKKDSLSEA ERAAAQAYAK EKGLTPPSTD
 61 HQDSGNTEAK GAEAIYNRVK AAKKVPLDRM PYNLQYTVEV KNGSLIIPHY DHYHNIKFEW
121 FDEGLYEAPK GYSLEDLLAT VKYYVEPRNA SDHVRKNKAD QDSKPDEDKE HDEVSEPTHP
181 ESDEKENHAG LNPSADNLYK PSTDTEETEE EAEDTTDEAE IPGTPSIRQN AMETLTGLKS
241 SLLLGTKDNN TISAEVDSLL ALLKESQPAP IQ   (SEQ ID NO : 258)
```

FIGURE 34